Calculating the expected automorphism group for RRTs

David Matthews

February 4, 2014

1 Notation

$\{T_i\}_{i=1}^n$ A random recursive tree process on n vertices.

$\mathcal{T}_n$ The set of of random recursive tree processes on n vertices.

$\tilde{\mathcal{T}}_n$ The set of labelled rooted tree on n vertices.

S_n the symmetric group on n elements.

$I(\sigma, T)$

$$\begin{cases} 1 & \text{if } \sigma \cdot T \in \mathcal{T}_n \\ 0 & \text{otherwise} \end{cases}$$

$P_n(\sigma)$ $\sum_{T \in T_n} I(\sigma, T)$

2 Random recursive trees

A *random recursive tree* (RRT) is a labelled, rooted tree obtained by assigning a root vertex and adding $n-1$ vertices one by one such that each new vertex is joined by an edge to a randomly and uniformly chosen existing vertex. It is natural to consider RRTs as nested sequences of rooted, labelled trees

$$T_1 \subset T_2 \subset \cdots \subset T_n$$

Where each T_t has precisely t vertices (and $(t-1)$ edges). At time t vertex v is chosen uniformly at random from $V(T_{t-1})$ and a new vertex v_t is attached to T_{t-1} via the edge $\{(v, v_t)\}$. Furthermore, we use the notation $\{T_i\}_{i=1}^n$ to mean a RRT on n vertice and we denote the set of all RRTs on n vertices by $\mathcal{T}$

Let $T = (V(T), E(T)$ be a labelled tree (not necessarily a RRT) and $d(v, w)$ be the length of the (unique) shortest path between any pair of vertices $v, w \in V(T)$. Every vertex $v \neq 1$ has a well defined *father*: the unique vertex v' adjacent to v such that $d(v', 1) < d(v, 1)$. By the process of RRT construction any vertex $1 \neq v \in T$ is adjacent to exactly one vertex with a lesser label.

Lemma 2.1. *Let $\mathcal{F}_n$ be the set of functions $f : N \longrightarrow N$ such that $f(1) = 1$ and $f(i) < i$ for $i = 2, 3, \ldots n$. There is a bijection between T_n and $\mathcal{F}$.*

Proof. One can associate a function $f \in \mathcal{F}$ to any RRT $\{T_i\}_{i=-1}^{i=n} \in \mathcal{T}_n$ by assigning $f(1) = 1$ and $f(i)$ the father of i. To see the converse, take any $f \in F_n$ and build $\{T_i\}_{i=1}^n$ by setting T_1 to be the graph with one vertex and no edges and subsequent T_i to be the graph built from $T_i - 1$ by attaching vertex i to $f(v)$ for $i = 2, 3, \ldots, n$. $\square$

Corollary 1. $|\mathcal{T}_n| = n - 1!$

Proof. Since $|\mathcal{T}_n| = |\mathcal{F}_n|$ it is enough to enumerate $\mathcal{F}_n$. One can write any $f \in \mathcal{F}_n$ as:

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & \ldots & n \\ 1 & f(2) & f(3) & f(4) & \ldots & f(n) \end{pmatrix}$$

Subject to $f(1) = 1$ and $f(i) < i$ for $i = 1, 2, \ldots n$. Note that f has 1 *choice* for $f(2)$ (i.e. $f(2) = 1$), two choices for $f(3)$ and, more generally, $i-1$ choices for $f(i-1)$. Therefore, $|\mathcal{F}_n| = n - 1!$ $\square$

Let $\tilde{\mathcal{T}}_n$ be the set of labelled rooted tree on n vertices. The symmetric group, S_n, can act on $\tilde{\mathcal{T}}_n$ by which permuting the non-root vertices of any $T \in \tilde{\mathcal{T}}_n$. Given a permutation $\sigma \in S_n$ and a tree $T \in \tilde{\mathcal{T}}_n$ we write the action of σ on T as $\sigma \cdot T$. Figure 1 shows that this action does not restrict to RRTs. This begs the question: Given $T \in T_n$ and $\sigma \in S_n$ under what conditions is $\sigma \cdot T \in T_n$?

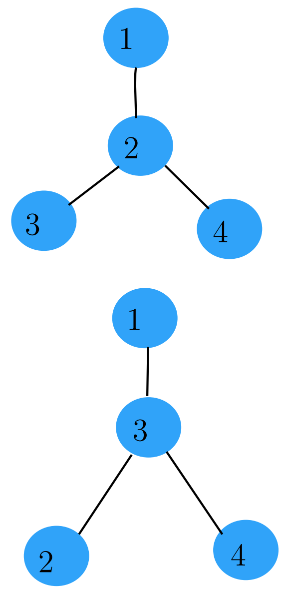

Figure 1: The top tree, T, is a RRT on n vertices. The bottom tree, $\sigma \cdot T$, where $\sigma = (23)$ and it is clear that $\sigma \cdot T \notin \mathcal{T}_4$,

Remark 1. *In fact no transposition σ of two non-root adjacent vertices i, j can be such that $\sigma \cdot T \in \mathcal{T}_n$. Without loss of generatility assume that $i < j$. Since i and j are adjacent $f(j) = i$, hence:*

$$\sigma(j) = i < j = \sigma(i) = \sigma(f(j))$$

Lemma 2.2. *Let $T \in T_n$ correspond to $f \in \mathcal{F}$ then $\sigma \cdot T$ corresponds to the following function:*

$$f' = \begin{pmatrix} 1 & \sigma(2) & \sigma(3) & \sigma(4) & \dots & \sigma(n) \\ 1 & \sigma(f(2)) & \sigma(f(3)) & \sigma(f(4)) & \dots & \sigma(f(n)) \end{pmatrix}$$

Proof. Let $T' = \sigma \cdot T$, there exists some function g corresponding to T' such that:

$$g = \begin{pmatrix} 1 & \sigma(2) & \sigma(3) & \sigma(4) & \dots & \sigma(n) \\ 1 & g(\sigma(2)) & g(\sigma(3)) & g(\sigma(4)) & \dots & g(\sigma(n)) \end{pmatrix}$$

Where $g(\sigma(i))$ is the father of $\sigma(i)$ but it is clear that the father of $\sigma(i)$ is $\sigma(f(i))$ hence $g(i) = \sigma(f(i))$ for $i = 2, 3, \dots, n$. □

Corollary 2. *Let $T \in \mathcal{T}_n$ and $\sigma \in S_n$. Then $\sigma \cdot T \in \mathcal{T}_n$ if and only if $\sigma(i) < \sigma(f(i))$.*

We define an indicator function for any $\sigma \in S_n, T \in \mathcal{T}_n$ as follows:

$$I(\sigma, T) = \begin{cases} 1 & \text{if } \sigma \cdot T \in \mathcal{T}_n \\ 0 & \text{otherwise} \end{cases}$$

2.1 Transpositions

In order to understand the effect of permutations of vertices on RRTs we shall examine $\sigma \cdot T$ where $\sigma = (i, j)$ is a transposition such that (without loss of generality) $i < j$.

By Corollary 2 if $\sigma \cdot T \in T_n$ the the corrosponding function, f satisfies that $\sigma(i) < \sigma(f(i))$ for $i = 2, 3, \ldots, n$.

Lemma 2.3. *Given a RRT $\{T = T_i\}_{i=1}^n$ and a transposition $\sigma = (p, q)$ the labelled tree $\sigma \cdot T$ is a RRT if and only if $f(q) < p$ and p is a leaf in T_q .*

Remark 2. *The idea behind the proof of this Lemma is that any $f \in \mathcal{F}_n$ can be split up into 5 parts as follows:*

$$f = \left(\begin{array}{ccc|c|ccc|c|ccc} 1 & \ldots & i-1 & i & i+1 & \ldots & j-1 & j & j+1 & \ldots & n \\ f(1) & \ldots & f(i-1) & f(i) & f(i+1) & \ldots & f(j-1) & f(j) & f(j+1) & \ldots & f(n) \end{array} \right)$$

We then notice that the first and fifth parts (with domain $i < p$ and $i > q$ respectively) are irrelevant to whether $\sigma \cdot T$ is a random recursive tree. It remains to find necessary and sufficient conditions for the second third and fourth parts such that $\sigma \cdot T \in \mathcal{T}_n$.

Proof. [of Lemma 2.3]

Case 1 $(i < p)$**.** If T is a RRT then $f(i) < p$ so $\sigma(i) = i$ and $\sigma(f(i)) = f(i)$. Hence, if T is a RRT, $\sigma(i) = i < f(i) = \sigma(f(i))$.

Case 2 $(i = p)$**.** If T is a RRT then $f(p) < p$ so $\sigma(f(p)) = f(p)$. Therefore, if T is a random recursive tree $\sigma(p) = q > p > f(p) = \sigma(f(p))$.

Case 3 $(p < i < q)$**.** Since $i \neq p$ and $i \neq q$, $\sigma(i) = i$. Also note that if T is a RRT $f(i) < i < q$.

If $f(i) = p$ then $\sigma(f(i)) = q > i = \sigma(i)$ hence T is *not* a RRT.

Case 4 $(i = q)$**.** If $f(q) = p$ then clearly $\sigma \cdot T$ is *not* an RRT since it (p, q) flips the edge (pq). If T is a RRT and $f(q) > p$ then $\sigma(f(q)) = f(q)$, therefore $\sigma(q) = p < f(q) = \sigma(f(q))$ so $\sigma \cdot T \notin \mathcal{T}_n$.

Case 5 $(i > q)$**.** Since $i \neq p$ and $i \neq q$ it is the case that $\sigma(i) = i$ hence

$$\sigma(f(i)) = \begin{cases} f(i) & \text{if } f(i) \neq p, q \\ p & \text{if } f(i) = q \\ q & \text{if } f(i) = p \end{cases}$$

Since $f(i), p, q < i$ it is always the case that $\sigma(i) < \sigma(f(i))$.

□

For a fixed $\sigma \in S_n$, we write $P_n(\sigma) = \sum_{T \in T_n} I(\sigma, T)$.

Lemma 2.4. $P_n(p, q) = \frac{(i-1)^2}{(j-1)(j-2)}$

Proof. Note that we can think of $P_n(p, q)$ as the number of trees $T \in \mathcal{T}_n$ such that $(p, q) \cdot T \in \mathcal{T}_n$. By Lemma 2.3, $\sigma = (p, q) \cdot T \in T_n$ if and only if p is a leaf in T_q and $f(q) < p$ so $P_n(p, q)$ is the number of trees $T \in \mathcal{T}_n$ such that p is a leaf in T_q and $f(q) < p$.

For every $T \in \mathcal{T}_n$ the associated function f can be split up into 5 parts as described in Remark 2, in particular the follwing matrix shows the number of possible values of $f(i)$ such that $\sigma \cdot T \in \mathcal{T}_n$ given each i:

$$f = \left(\begin{array}{cccc|c|ccc|c|ccc} 1 & 2 & 3 & \ldots & i & i+1 & \ldots & j-1 & j & j+1 & \ldots & n \\ 1 & 1 & 2 & \ldots & i-1 & i-1 & \ldots & j-2 & i-1 & j & \ldots & n-1 \end{array} \right)$$

Therefore,

$$P_n(p, q) = \frac{(p-1)^2}{(q-1)(q-2)}(n-1)!$$

□